

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平10-65733

(43)公開日 平成10年(1998)3月6日

(51)Int.Cl. <sup>6</sup>	識別記号	序内整理番号	F I	技術表示箇所
H 04 L 12/56		9744-5K	H 04 L 11/20	1 0 2 D
	12/28		H 04 M 3/00	D
H 04 M 3/00			H 04 Q 3/00	
H 04 Q 3/00			H 04 L 11/00	3 1 0 D

審査請求 有 請求項の数4 OL (全 15 頁)

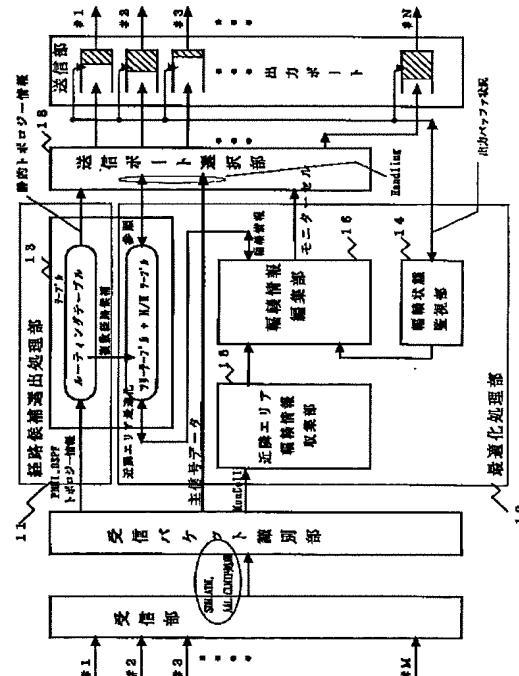
(21)出願番号	特願平8-215678	(71)出願人	000004237 日本電気株式会社 東京都港区芝五丁目7番1号
(22)出願日	平成8年(1996)8月15日	(72)発明者	升田 道雄 東京都港区芝五丁目7番1号 日本電気株式会社内
		(72)発明者	西原 基夫 東京都港区芝五丁目7番1号 日本電気株式会社内
		(72)発明者	小川 亮 東京都港区芝五丁目7番1号 日本電気株式会社内
		(74)代理人	弁理士 山川 政樹

(54)【発明の名称】 高速ルーティング制御方式

(57)【要約】

【課題】 宛先ノードが近接する場合でも最適な経路選択を実施でき、かつ迂回路への切替処理が高速化できるとともに、トポロジー情報によるトラフィック量を抑制でき、ネットワークの大規模化に追従できる高速ルーティング制御方式を提供することを目的とするものである。

【解決手段】 経路候補選出処理部11にて、各ノード間を接続するリンクの物理的な接続情報に基づいて宛先ノードまでの経路候補を複数選択し、最適化処理部12にて、近隣エリアに属する各リンクの輻輳状態に基づいて、経路候補選出処理部11により選出された複数の経路候補のうち、輻輳が発生しているリンクを含まない経路候補を最適経路として特定する。



**【特許請求の範囲】**

**【請求項1】** コネクション型のATM網を介して LANなどのコネクションレスデータからなるパケットを所定の宛先ノードに対して転送する場合のパケット転送経路を選択するルーティング制御方式において、各ノード間を接続するリンクの物理的な接続情報に基づいて宛先ノードまでの経路候補を複数選択する経路候補選出処理部と、所定の近隣エリアに属する各リンクの輻輳状態に基づいて、経路候補選出処理部により選出された複数の経路候補のうち、輻輳が発生しているリンクを含まない経路候補を最適経路として特定する最適化処理部とを備えることを特徴とする高速ルーティング制御方式。

**【請求項2】** 請求項1記載の高速ルーティング制御方式において、

経路候補選出処理部は、予め設定された各リンクの選出優先度を示すリンクコスト値に基づいて、宛先ノードへの経路のうち各経路を構成するリンクのリンクコスト値の総和が最も小さい最小コスト経路を選出するとともに、自己ノードに隣接する個々の隣接ノードを経由する宛先ノードへの経路のうち、その経路のリンクコスト値の総和が、前記最小コスト経路のリンクコスト値の総和と自ノードから前記隣接ノードまでのリンクコスト値の2倍との和より小さい経路を選出するようにしたことを特徴とする高速ルーティング制御方式。

**【請求項3】** 請求項1記載の高速ルーティング制御方式において、

最適化処理部は、隣接ノードから所定のモニターセルを受信し、そのモニターセルに格納されている輻輳情報のうち、自ノードから所定リンク数以上離れたノードが収容するリンクの輻輳情報を廃棄するとともに、残りの各ノードが収容する各リンクの輻輳情報を取得して最適経路の特定に用い、取得した輻輳情報と自ノードが収容する各リンクの輻輳情報とから新たなモニターセルを生成して隣接ノードに送信するようにしたことを特徴とする高速ルーティング制御方式。

**【請求項4】** 請求項3記載の高速ルーティング制御方式において、

最適化処理部は、所定周期ごとにモニターセルを送信するようにしたことを特徴とする高速ルーティング制御方式。

**【発明の詳細な説明】**

**【0001】**

**【発明の属する技術分野】** 本発明は、高速ルーティング制御方式に関し、特にATM網で、ATMレイヤのVPIおよびVCIにより、IPパケットをATMレイヤの情報のみで識別し、パケット組立処理遅延なしに転送処理を行うATMコネクションレス通信装置の高速ルーティング制御方式に関するものである。

【0002】

**【従来の技術】** 近年、ユーザ側LAN (Local Area Network) のローカルなコネクションレス特性を、グローバルなATM網で収容して、高速かつ低価格で運用する要求が高まっている。本来のATMネットワークでは、シグナリングなどにより、主信号データ転送に先立ってコネクション設定を行い、コネクション設定ごとに各スイッチで詳細なコスト評価、CAC (Connection Admission Control) 計算などを行い、複数の候補から最適な経路を選択すると同時に、ユーザからのQoS (Quality of Service) 要求を満たすためのリソース確保を行うものとなっている。

**【0003】** ATM網におけるPNNI (Private Network-to-Network Interface) およびIP通信におけるOSPF (Open Shortest Path First) で代表されるルーティングプロトコルでは、最適な経路を選択するために、「Dijkstra」アルゴリズムを用いている。

「Dijkstra」アルゴリズムは、全てのノードからのリンク情報を元に同一コストの最小パスを計算し、自分をルート (root: 根) とする各宛先までのツリー (ルートマップ: 最小木) を作成するものであり、「Dijkstra」アルゴリズムを実行するために各ノードは、経路評価として直に接続するリンク状態をなわちトポロジーを管理し、その情報を全ノードに向けて通知することが必要となる。

**【0004】** そこで、PNNIやOSPFでは、リンク情報を通知する方法として「Flooding」と呼ばれる手法を規定している。この方法は、トポロジー情報を含むパケットを一定エリア内に早急に伝搬するために用いられ、受信した情報が自分の持っている情報より新しければ周りの皆に配布し、古ければ廃棄する、という動作が基本になっている。図2は、PNNIでの「Flooding」によるトポロジー情報の配布方法を示した概念図である。A.1-X.1間リンク状態が変化した場合に、その変化を検出したノードA.1が発信源となって特定の論理的なエリアPG (A) 内に「洪水」のように広告することを特徴とする伝搬方式である。

**【0005】** 一方、IPパケット通信に代表される通常のコネクションレス網の通信方式においては、ユーザからのQoS要求がなく、ネットワーク側でシグナリング処理により経路を評価するという2次的な最適化処理を行わない。従来、この種のコネクションレス通信方式は、回線断および瞬間的な一方路への集中に対しても、パケットが廃棄されることなく最大限に通信要求を満足できることを目的として、直通経路以外に迂回経路を設定しておき、動的に迂回経路への切り替え処理が可能となる手法が用いられている。

**【0006】** ATM網でコネクションレス通信を行うにあたり、迂回パスへの選択機能を具備した方式として、

複数の迂回経路となる複数候補を用意し、候補経路ごとに余剰帯域値を監視し、余剰帯域のより大きい経路へ切り替える制御を行う方法が提案されている（例えば、特開平3-242063号公報など）。図3は、ATM網における従来のコネクションレス処理装置の一例を示すブロック図であり、受信部31、相手アドレス認識部32、記憶テーブル部33、送信制御部34、送信部35から構成される。

**【0007】**図4は、従来の送信制御部の動作フローを示したフローチャートである。同図に示すように、従来の方法は、コネクションレスパケットを受信すると、パケットの付加情報から宛先アドレスを抽出し、宛先アドレスに対応したPVC(Permanent Virtual Connection)を選択する。迂回経路設定手段として、図3のテーブル記憶部33で直通経路および複数の迂回経路のPVCの使用量を管理し、送信制御部34で、最も空きのあるPVCを迂回経路として選択することによりユーザの通信要求を満足させようとするものである。

#### 【0008】

【発明が解決しようとする課題】しかしながら、このような従来のルーティング制御方式では、次のような問題点があった。第1の問題点は、ルーティング処理が高速化できない、ということである。すなわち、「Dijkstrijkstra」アルゴリズムに代表されるリンク状態型の経路評価による経路の確定は、伝搬方式としてブロードキャスト型の「Flooding」を用いることにより、ルーティング情報交換に要するトラヒック量が増大するため、頻繁な情報交換が行われない。したがって、リアルタイム性に限界があり、ルーティング時の最適性が粗くなる。

**【0009】**その精査手段として、ATM網におけるPNNIシグナリングでは、コネクション設定ごとにCAC処理やクランクバック処理などにより2次的に詳細で、かつユーザーが申告するQoSを満足するような経路評価を行い、ルーティング経路を確定する。しかし、このシグナリング処理のオーバーヘッドにより遅延が発生する。従来の迂回経路選択方法では、各経路の許容帯域値の精度を上げるために、トラヒック変動特性を予測し、実リンクトポロジー変動に忠実なコスト関数を求めるなどの複雑な計算が必要となり、迂回経路へ切り替えるための判定処理が高速化できない。問題となる要因は、いずれの方法も規定時間内に最適解となる迂回経路を導き出すための高速、低遅延を優先した処理手段が設けられていないことがある。

**【0010】**第2の問題点は、トポロジー情報伝搬においてトラヒック量が増大する、ということである。従来のコスト評価の判断基準としている各経路の許容帯域は、各コネクションレス処理装置のバッファ占有率、強いてはトラヒック特性に依存し、刻々と変動する。したがって、各経路の許容帯域の変動特性に引きづられて、

ルーティングテーブルが更新され、ネットワーク内の経路制御が混乱状態に陥る。また、トポロジー変動を各コネクションレス通信装置に広告（アドバタイズ）するためのトラヒック量が増大する。問題となる要因は、コスト情報に余剰帯域値を反映していることと、伝搬方式として「Flooding」を用いていることにある。

**【0011】**第3の問題点は、ネットワークの大規模化に追随しにくい、ということである。問題となる直接要因は、前述と同様、伝搬方式が「Flooding」型であることに起因し、コネクションレス通信装置数をNとすると、 $O(N^2)$ のオーダーでトポロジー情報伝搬トラヒック量が増加する。トラヒック量を抑制する手段としては、下記3点が考えられている。第1の手段は、各コネクションレス通信装置でトラヒック量を抑制するタイマを配備する方法である。第2の手段は、トポロジー情報交換を行うパケットのヘッダ部の、寿命（Lifetime）を使用し、規定時間になればパケットを死滅させる方法である。

**【0012】**第3の手段は、「Flooding」する論理エリアを分割階層化して、広告範囲を限定し、エリア間および下位エリア、上位エリア間は要約した情報を伝達する方法である。図2に、階層化されたネットワークにおける「Flooding」方法を示す。下位層のPG(A)-PG(X)間にはA.1-X.1リンクとA.2-X.2リンクの2本のリンクが存在するが、上位層では、2本のリンクをA-X間の1本のリンクに集約したトポロジー情報（Topology Aggregation）を「Flooding」していることを示している。

**【0013】**このような抑制手段により、過度のトラヒック増大現象をある程度緩和できるが、分割エリア数をmとしたとき、トポロジー情報伝搬トラヒック量は $O(N^2/m)$ のオーダーで依然として増大傾向にあり、ネットワーク大規模化に追随しにくい。また、分割エリア数(m)を増やしてトポロジー情報の広告範囲を縮小した場合は、トポロジー情報集約時に誤差が生じるため、結果としてルーティングの正確性を失うという問題点があった。本発明はこのような課題を解決するためのものであり、近接する場合でも最適な経路選択を実施でき、かつ迂回路への切替処理が高速化できるとともに、トポロジー情報によるトラフィック量を抑制でき、ネットワークの大規模化に追従できる高速ルーティング制御方式を提供することを目的とするものである。

#### 【0014】

【課題を解決するための手段】このような目的を達成するために、本発明による高速ルーティング制御方式は、各ノード間を接続するリンクの物理的な接続情報に基づいて宛先ノードまでの経路候補を複数選択する経路候補選出処理部と、所定の近隣エリアに属する各リンクの輻輳状態に基づいて、経路候補選出処理部により選出された複数の経路候補のうち、輻輳が発生しているリンクを

含まない経路候補を最適経路として特定する最適化処理部とを備えるものである。したがって、各ノード間を接続するリンクの物理的な接続情報に基づいて宛先ノードまでの経路候補が複数選択され、これら経路候補のうち、所定の近傍エリアに属する各リンクの輻輳状態に基づいて、輻輳が発生しているリンクを含まない経路候補が最適経路として特定される。また、経路候補選出処理部は、予め設定された各リンクの選出優先度を示すリンクコスト値に基づいて、宛先ノードへの経路のうち各経路を構成するリンクのリンクコスト値の総和が最も小さい最小コスト経路を選出するとともに、自己ノードに隣接する個々の隣接ノードを経由する宛先ノードへの経路のうち、その経路のリンクコスト値の総和が、最小コスト経路のリンクコスト値の総和と自ノードから隣接ノードまでのリンクコスト値の2倍との和より小さい経路を選出するようにしたものである。したがって、予め設定された各リンクのリンクコスト値に基づいて、宛先ノードへの経路のうちリンクコスト値の総和が最も小さい最小コスト経路が選出されるとともに、各経路のうち、リンクコスト値の総和が、最小コスト経路のリンクコスト値の総和と自ノードから隣接ノードまでのリンクコスト値の2倍との和より小さい経路が選出される。

【0015】また、最適化処理部は、隣接ノードから所定のモニターセルを受信し、そのモニターセルに格納されている輻輳情報のうち、自ノードから所定リンク数以上離れたノードが収容するリンクの輻輳情報を廃棄するとともに、残りの各ノードが収容する各リンクの輻輳情報を取得して最適経路の特定に用い、取得した輻輳情報と自ノードが収容する各リンクの輻輳情報とから新たなモニターセルを生成して隣接ノードに送信するようにしたものである。したがって、受信したモニターセルに格納されている輻輳情報のうち、自ノードから所定リンク数以上離れたノードが収容するリンクの輻輳情報が廃棄されるとともに、残りの各ノードが収容する各リンクの輻輳情報が取得されて最適経路の特定に用いられ、取得した輻輳情報と自ノードが収容する各リンクの輻輳情報とから新たなモニターセルが生成されて隣接ノードに送信される。また、最適化処理部は、所定周期ごとにモニターセルを送信するようにしたものである。したがって、所定周期ごとにモニターセルが送信される。

#### 【0016】

【発明の実施の形態】次に、本発明について図面を参照して説明する。図1は本発明の一実施の形態であるコネクションレス通信装置の高速ルーティング制御方式を示すブロック図であり、基本ルーティング処理(Basic Routing)部11と最適化処理(Optimize)部12からなる2段階のルーティング制御方式により構成されていることを特徴としている。経路候補選出処理部11は、各ノードの物理的な接続情報が予め設定されているルーティングテーブル11aを有し、ATM網で利用されるP

NNI(Private Network-to-Network Interface)およびIP通信におけるOSPF(Open Shortest Path First)で代表されるルーティングプロトコルなどで用いられているリンクステート型の経路評価アルゴリズムに基づいて、大局的なルーティング経路候補を複数選出する機能を有している。

【0017】最適化処理部12は、経路候補選出処理部11にて選出された複数のルーティング経路候補の中で、さらに詳細な経路評価を行い、特定の近傍エリア内でリアルタイムに最適経路を選出する機能を有しており、図1に示すように、テーブル部13、輻輳状態監視部14、近傍エリア輻輳情報収集部15、輻輳情報編集部16から構成されている。輻輳状態監視部14では、自身のバッファのキュー長またはキュー増加率などを監視して、出力ポート単位に輻輳状態であるか否かを検出する。

【0018】近傍エリア情報収集部15は、図5に示すように、輻輳状況モニターセルを送出する送出部54、輻輳状況モニターセルを受信する受信部53、受信した輻輳状況モニターセルの距離すなわち経由リンク数(Hop数)を加味してエージング(Aging:寿命)判定を行うエージング判定部52から構成されている。また、近傍エリア情報収集部15は、タイマ51を具備しており、隣接するコネクションレス通信装置に対して、自己タイミングで輻輳状況モニターセルを定期的に生成して送出する機能を有している。

【0019】したがって、近傍エリア情報収集部15は、隣接するコネクションレス通信装置から輻輳状況を格納したモニターセルを受信部53で受信し、その輻輳状況に変化があった場合には、隣接するコネクションレス通信装置とのリンクを、後述するテーブル13のツリーテーブルから除いた経路候補を検索して準最適経路を選出する。さらに、ツリーテーブルを参照して、経路候補が変更されたことをH/Wテーブルに反映する。

【0020】各コネクションレス通信装置が収集するモニターセルには、近傍コネクションレス通信装置の出力ポート単位にバッファの輻輳状況を示す情報とともに、到着するまでに経由してきたコネクションレス通信装置のHop数を示すフィールドが設けられており、経由するコネクションレス通信装置のHop数が規定回数に達した場合は、該当情報を廃棄することにより近傍エリア外への広告を抑制するエージング判定に用いられる。エージング判定部52では、管理するHop数の制限値(a)と比較して、自分が獲得し他の隣接ノードに配布する必要性を判断し、保持または廃棄のいずれかを選択して処理するものとなる。

【0021】輻輳情報編集部16では、自身のテーブル13のツリーテーブルを参照し、配布するモニターセルの編集を行う機能を有する。図6はモニターセルフォーマットであり、同図に示すように、自身の輻輳状況を先

頭に記述し、その後に続けて近隣エリア情報収集部におけるエージング判定処理後の情報を挿入したフォーマットでモニターセルを生成する。コネクションレス通信装置からの距離（Hop数）に応じて輻輳情報が記述される点が特徴である。

【0022】次に、図7を参照して、本発明の動作について説明する。図7は、本発明を適用した高速ルーティング制御方式を示す動作概念図であり、(a)はコネクションレス通信装置の物理的トポロジー、(b)は経路候補選出処理部11の動作、(c)は最適化処理部12の動作を示している。以下では、A局-J局間で通信を行う場合のルーティング制御動作を例として本発明の動作を説明する。なお、ネットワーク構成時には、各リンク速度の逆数、遅延特性、あるいは廃棄率などが評価され、各リンクの選出優先度を示すリンクコスト値が各リンクに予め割り当てられているものとする。図7(b)の()内数字は各リンクのリンクコスト値を示している。

【0023】まず、各コネクションレス通信装置の経路候補選出処理部11は、ルーティングテーブル11aを参照して各リンクのリンクコスト値を獲得し、これらリンクコスト値を基に「Dijkstra」アルゴリズムを用いて、自分A局をルート（root：根）とした宛先J局までのルートマップ（最小木）を作成する。さらに、迂回経路対象候補となる複数経路を算出する。ここでいう複数経路とは、最小コスト経路のリンクコスト値の総和と同値、あるいは後述する「Loop free」となる条件を加味して選出したものである。この処理により選出したA-J局間の複数経路が図7(b)に実線で示されている。

【0024】最適化処理部12では、経路候補選出処理部11より提示された複数の候補経路の中で、自分の近傍エリア内の輻輳状況を加味した最適化を行い、ルーティング経路を特定する。この処理により選出された最適経路が図7(c)に実線で示されている。ここで、図9を参照して、経路候補選択部11aでの複数候補選出過程について説明する。図9は複数経路選出過程を示す説明図である。

【0025】候補となる複数経路は、基本ルーティング部11で選ばれ、かつ重複経路すなわちルーピング（Looping）が発生しない経路でなければならない。今、A局からB局までの各リンクのリンクコスト値の総和をNAとし、すべてのノード間リンクコストをc（一定）とすると、最小リンクコスト値の総和は、 $NA = 4 \times c$ となる。また、A局に隣接してa1～a6局が存在し、A-a<sub>n</sub>間（n:1～6）の最小リンクコスト値はcである。

【0026】a<sub>n</sub>-B間のリンクコスト値をN<sub>a<sub>n</sub></sub>（n:1～6）とし、AからBへ転送する場合において、a<sub>n</sub>を経由してもルーピングしないためには、A局

に戻らない条件を考慮すると、

$$Na_n < NA + c = 5c \quad (\text{式1})$$

となる。このとき、A局からみたa<sub>n</sub>局経由のB局までのリンクコスト値は、

$$Na_n + c < NA + c + c = 6c \quad (\text{式2})$$

となる。

【0027】したがって、複数経路の対象となる候補は、同じ最小リンクコスト値Nを持つ経路すなわち最小コスト経路と、隣接する局までの最小リンクコスト値を

cとした時、その局を経由するA-B間の最小リンクコスト値が、N+2×cよりも小さくなる経路、すなわち「loop free」の2種類である。図9の例では、A-B間の複数候補選択過程で、A局から次に選択し得るノードはa1, a2, a4, a6の4つの隣接ノードに限定される。A局がa3またはa5を経由してB局に到着する経路は最小でもリンクコスト値が「6c」となるため、(式2)より、A-B間経路候補から除外されることを示している。

【0028】また、図8は基本ルーティング部11で選出された経路候補がツリーテーブル化される過程と、送信ポート選択部18で実際のルーティングに使用されるH/Wテーブルがツリーテーブルを参照して準最適経路を特定する過程を示したものである。例えば、図7(c)のArea2内に属するE局に着目した場合、作成されるツリーテーブル81は、自分(E)をルートにした他の全ノードを宛先とする最適経路をCI(Congestion Indication: 輻輳表示)フラグ付きで出力ポート単位に管理する。

【0029】図8の例では、「候補1」の最適パス以外に、G-HリンクおよびH-Jリンクをそれぞれ除外（ブロック）して「Dijksta」演算を行った結果、得られた経路をそれぞれ「候補2」および「候補3」として選出したことを示している。CI情報はモニターセルによって収集され、準最適経路の特定は、ツリーテーブル81のCIフラグを参照して輻輳状況のチェックが行われた後、H/Wテーブル82に反映される。

【0030】ここで、ノードEからノードJへの経路のうち、H-Jリンクの出力ポートのバッファが輻輳を起こしていることが判明した場合、ツリーテーブル81上の該当箇所のCIフラグが”Nor（輻輳解除状態）”から”ALM（輻輳発生状態）”に更新される。したがって、実際のルーティング時は「候補3」のG-I-Jの経路を選択するものとなる。なお、H/Wテーブル82の作成手順については、まず、基本ルーティング部11より提示されたツリーテーブル81のリーフ（leaf）の位置に該当する宛先アドレスをサーチし（ステップ①）、該宛先アドレスに対応した次段転送先ポートを決定する。その際、ツリーテーブル81より提示された候補順に経路情報の輻輳状況の確認を行い（ステップ②）、最適経路を特定する（ステップ③）。

【0031】したがって、図1の送信ポート選択部18では、H/Wテーブル82を参照して、近傍エリア内のコネクションレス通信装置にバッファ溢れの危険がない定常時は、第1の最適経路を選択する。一方、近隣コネクションレス通信装置に局所的なバッファ溢れが発生していると判れば、与えられた複数候補の中から輻輳発生経路を避けた任意(第2～第n)の出力方路を特定し、パケット単位に一時的に迂回させる制御を行う。

【0032】このように、経路候補選出処理部11にて、各ノード間を接続するリンクの物理的な接続情報に基づいて宛先ノードまでの経路候補を複数選択し、最適化処理部12にて、近隣エリアに属する各リンクの輻輳状態に基づいて、経路候補選出処理部11により選出された複数の経路候補のうち、輻輳が発生しているリンクを含まない経路候補を最適経路として特定するようになので、経路候補選択時には変動の少ない静的な情報に基づいて経路候補が選出されるため、経路候補の選出に要する処理時間を短縮できるとともに、これら経路候補のうち輻輳発生リンクの有無に基づいて最適経路が特定されるため、最適経路の特定に要する処理時間を短縮でき、全体としてルーティング制御を高速に処理することができる。

【0033】また、経路候補選出処理部は、予め設定された各リンクのリンクコスト値に基づいて、宛先ノードへの経路のうち各経路を構成するリンクのリンクコスト値の総和が最も小さい最小コスト経路を選出するとともに、自ノードに隣接する個々の隣接ノードを経由する宛先ノードへの経路のうち、その経路のリンクコスト値の総和が、最小コスト経路のリンクコスト値の総和と自ノードから隣接ノードまでのリンクコスト値の2倍との和より小さい経路を選出するようにしたので、比較的簡単な処理にて、自ノードを重複して経由する経路を事前に選別することが可能となり、宛先ノードが比較的近接する場合でも最適な経路選択を実施することができる。

【0034】次に、図10を参照して、最適化処理部12の動作について説明する。図10は、本発明を適用した最適化処理部12における輻輳状況伝搬方式の特徴を示した動作概念図である。同図に示すような格子型のネットワークを構成するノードGに着目する。ノードGは、まずノードIに対して、自己タイミングにより輻輳状況モニターセルを送出する。ノードGからノードIへの輻輳状況モニターセルには、前述(図6参照)のように、自身の現在の輻輳状況を先頭に、ノードGが現在獲得している隣接ノードD, E, Jの輻輳状況を付加して送出する。

【0035】さらに、ノードGは、他の隣接ノードであるノードD, E, Jに対しても輻輳状況モニターセルを同様に送出し、結果的にノードGの輻輳状況の広告範囲は、ノードGを中心として距離a(本例ではAge制限ホップ数=2)のエリア内に閉じられている。ノードG

と隣接関係にあるノードIもまた、自身のタイミングにより、現在の自己輻輳状況を先頭に、現在獲得している隣接ノードF, L, Kの輻輳状況を付加して、ノードGにモニターセルを送信する。他の隣接ノードであるノードD, E, Jも同様に動作し、ノードGは自身を中心とした距離aエリア内に存在するノードの輻輳状況を獲得する。

【0036】本発明を適用した最適化処理部における輻輳状況伝搬方式によれば、コネクションレス通信装置数の増加に比例して輻輳情報伝搬トラヒック量が増加する: O(N)のものとなり、トラヒック量の増加の割合が一定に保たれる。

【0037】図11は、図10の格子型ネットワークにおけるノードGの輻輳状況モニターセル送出タイミングを示したタイムチャートである。輻輳状況モニターセルを送出するタイミングは各ノードが一定周期 $\alpha$ (μsec)で行うことを示している。周期は各コネクションレス通信装置内のクロックに同期して行われる。この場合には、前述(図5, 6)したように、近隣エリア輻輳状態収集部15により、ある周期タイミングでノードI, D, Jから受信された輻輳状況モニターセル111～113に基づいてツリーテーブル81の各経路の輻輳情報が更新される。

【0038】その後、これら輻輳情報に基づいてノードE宛の輻輳状況モニターセル114が生成され次の周期タイミングに送信される。したがって、輻輳状況通知セル内の情報は、情報源となるコネクションレス通信装置の距離に比例して古くなり、最も遠いところでも、"距離(a) × 周期(α)"分の過去の情報を獲得できる。

近接ノードに対する輻輳情報ほど最近となり、獲得する情報としては適切であることが判る。

【0039】このように、最適化処理部12にて、隣接ノードから所定のモニターセルを受信し、そのモニターセルに格納されている輻輳情報のうち、自ノードから所定リンク数以上離れたノードが収容するリンクの輻輳情報を廃棄するとともに、残りの各ノードが収容する各リンクの輻輳情報を取得して最適経路の特定に用い、取得した輻輳情報と自ノードが収容する各リンクの輻輳情報とから新たなモニターセルを生成して隣接ノードに送信するようになので、モニターセルの広告範囲を物理的な近隣エリア内に限定することが可能となり、ノード数の増加に対して、輻輳情報を収集するためのモニターセルによるトラフィック増加の割合が常に一定となり、ネットワークの大規模化にも対応することができる。

【0040】また、所定周期ごとにモニターセルを送信するようにしたので、モニターセルのセル遅延揺らぎを抑制でき、各ノードが有するリンクバッファのバッファ占有率の変動を低減することができる。さらに、一部のノードから誤って多量のモニターセルが送信された場合でも、隣接するノードからは正常な周期で送信されるも

のとなり、ネットワーク全体へのモニターセルトラフィックの氾濫を防止することができる。

【0041】

【発明の効果】以上説明したように、本発明は、大局的なルーティングを行う経路候補選択処理部と、所定の近隣エリアにおける最適経路を特定する最適化処理部とを設けて、経路候補選択処理部にて、各ノード間を接続するリンクの物理的な接続情報に基づいて宛先ノードまでの経路候補を複数選択し、最適化処理部にて、これら経路候補のうち、所定の近隣エリアに属する各リンクの輻輳状態に基づいて、輻輳が発生しているリンクを含まない経路候補を最適経路として特定するようにしたので、経路候補選択時には変動の少ない静的な情報に基づいて経路候補が選出されるため、経路候補の選出に要する処理時間を短縮できるとともに、これら経路候補のうち輻輳発生リンクの有無に基づいて最適経路が特定されるため、最適経路の特定に要する処理時間を短縮でき、全体としてルーティング制御を高速に処理することができる。

【0042】また、経路候補選出処理部では、予め設定された各リンクの選出優先度を示すリンクコスト値に基づいて、宛先ノードへの経路のうち各経路を構成するリンクのリンクコスト値の総和が最も小さい最小コスト経路を選出するとともに、自ノードに隣接する個々の隣接ノードを経由する宛先ノードへの経路のうち、その経路のリンクコスト値の総和が、最小コスト経路のリンクコスト値の総和と自ノードから隣接ノードまでのリンクコスト値の2倍との和より小さい経路を選出するようにしたので、比較的簡単な処理にて、自ノードを重複して経由する経路を事前に選別することが可能となり、宛先ノードが比較的近接する場合でも最適な経路選択を実施することができる。

【0043】また、最適化処理部12では、隣接ノードから所定のモニターセルを受信し、そのモニターセルに格納されている輻輳情報のうち、自ノードから所定リンク数以上離れたノードが収容するリンクの輻輳情報を廃棄するとともに、残りの各ノードが収容する各リンクの輻輳情報を取得して最適経路の特定に用い、取得した輻輳情報と自ノードが収容する各リンクの輻輳情報とから新たなモニターセルを生成して隣接ノードに送信するようにしたので、モニターセルの広告範囲を物理的な近隣

エリア内に限定することが可能となり、ノード数の増加に対して、輻輳情報を収集するためのモニターセルによるトラフィック増加の割合が常に一定となり、ネットワークの大規模化にも対応することができる。また、所定周期ごとにモニターセルを送信するようにしたので、モニターセルのセル遅延揺らぎを抑制でき、各ノードが有するリンクバッファのバッファ占有率の変動を低減することができる。さらに、一部のノードから誤って多量のモニターセルが送信された場合でも、隣接するノードからは正常な周期で送信されるものとなり、ネットワーク全体へのモニターセルトラフィックの氾濫を防止することができる。

【図面の簡単な説明】

【図1】 本発明の一実施の形態による高速ルーティング制御方式を示すブロック図である。

【図2】 従来のトポロジー情報の配布方法を示す概念図である。

【図3】 従来のコネクションレス処理装置の一例を示すブロック図である。

【図4】 従来のコネクションレス処理装置の動作を示すフローチャートである。

【図5】 近隣エリア輻輳状態収集部を示すブロック図である。

【図6】 輻輳状態モニターセルのフォーマット例を示す説明図である。

【図7】 本発明による高速ルーティング制御方式を示す動作概念図である。

【図8】 ツリーテーブルおよびH/Wテーブルの作成過程を示す説明図である。

【図9】 複数経路選出過程を示す説明図である。

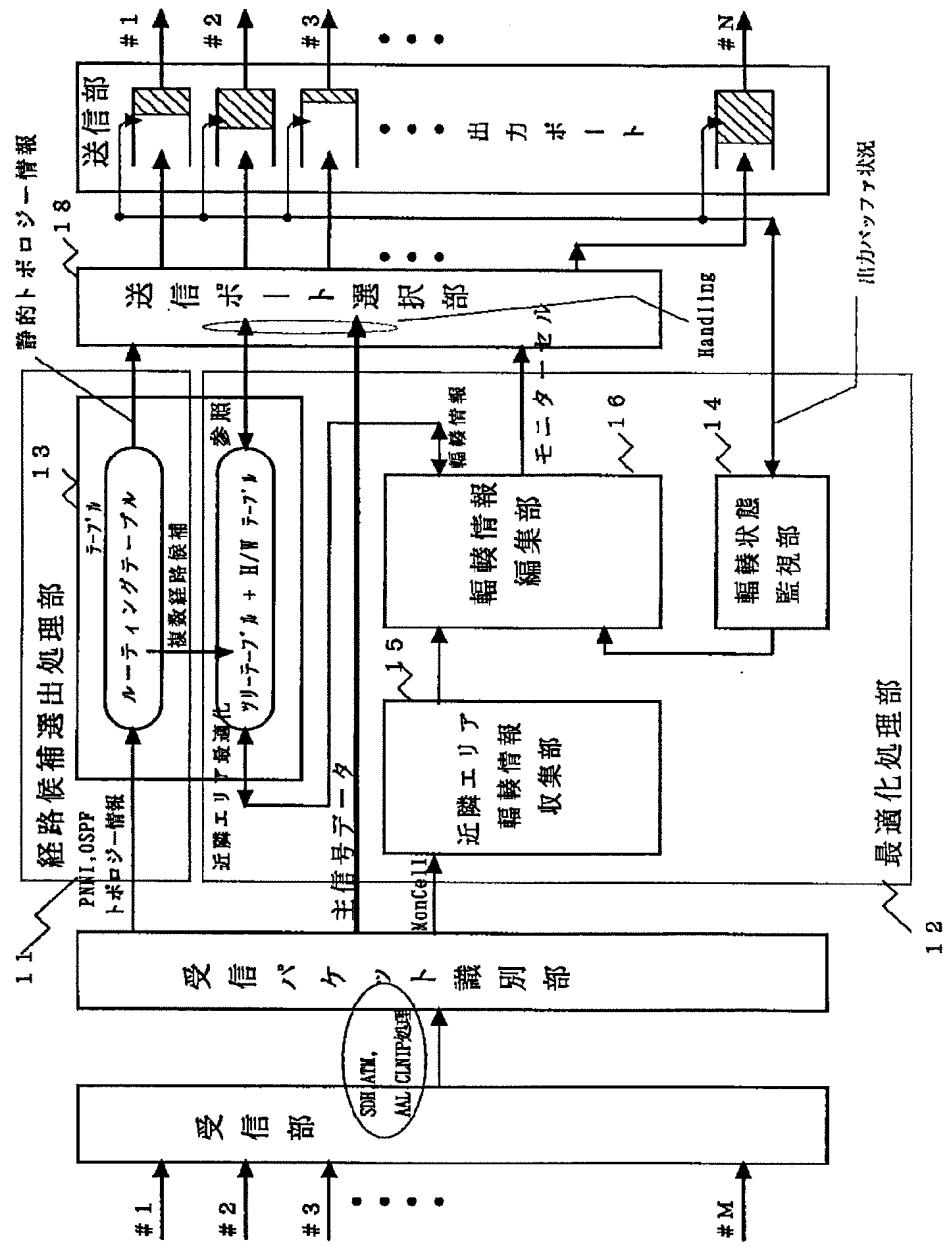
【図10】 最適化処理部の輻輳状況伝搬方式を示す動作概念図である。

【図11】 輻輳状況モニターセル送出タイミングを示すタイムチャートである。

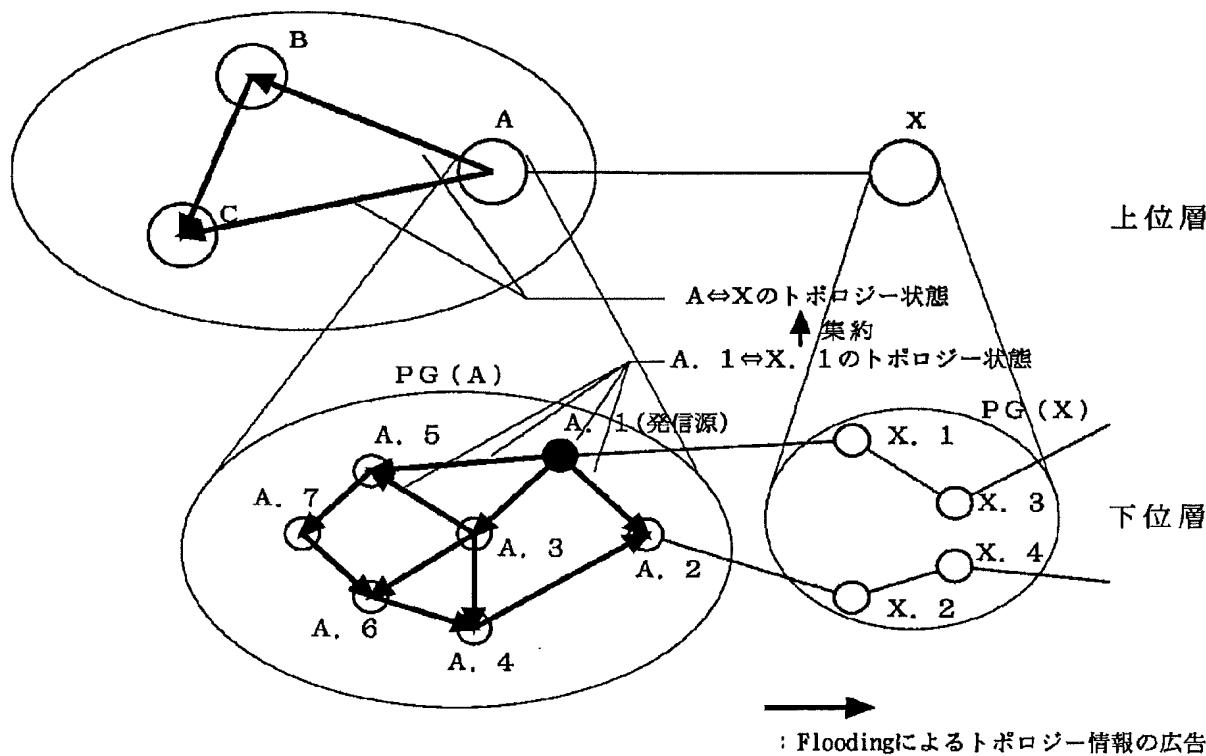
【符号の説明】

1 1…経路候補選択処理部、1 2…最適化処理部、1 3…テーブル部、1 4…輻輳状況監視部、1 5…近隣エリア輻輳情報収集部、1 6…輻輳情報編集部、5 1…タイマ、5 2…エージング判定部、5 3…受信部、5 4…送信部、8 1…ツリーテーブル、8 2…H/Wテーブル。

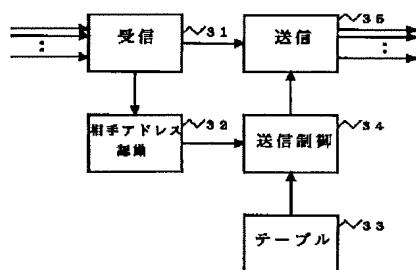
【図1】



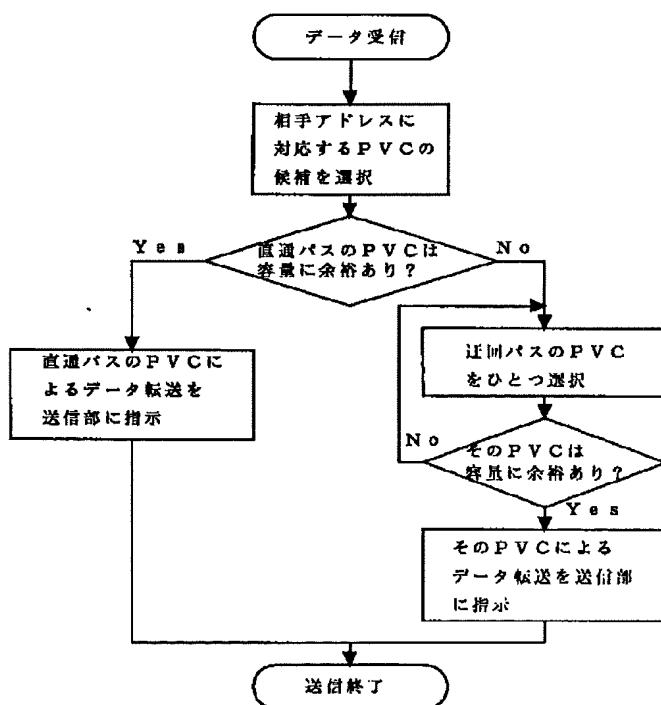
【図2】



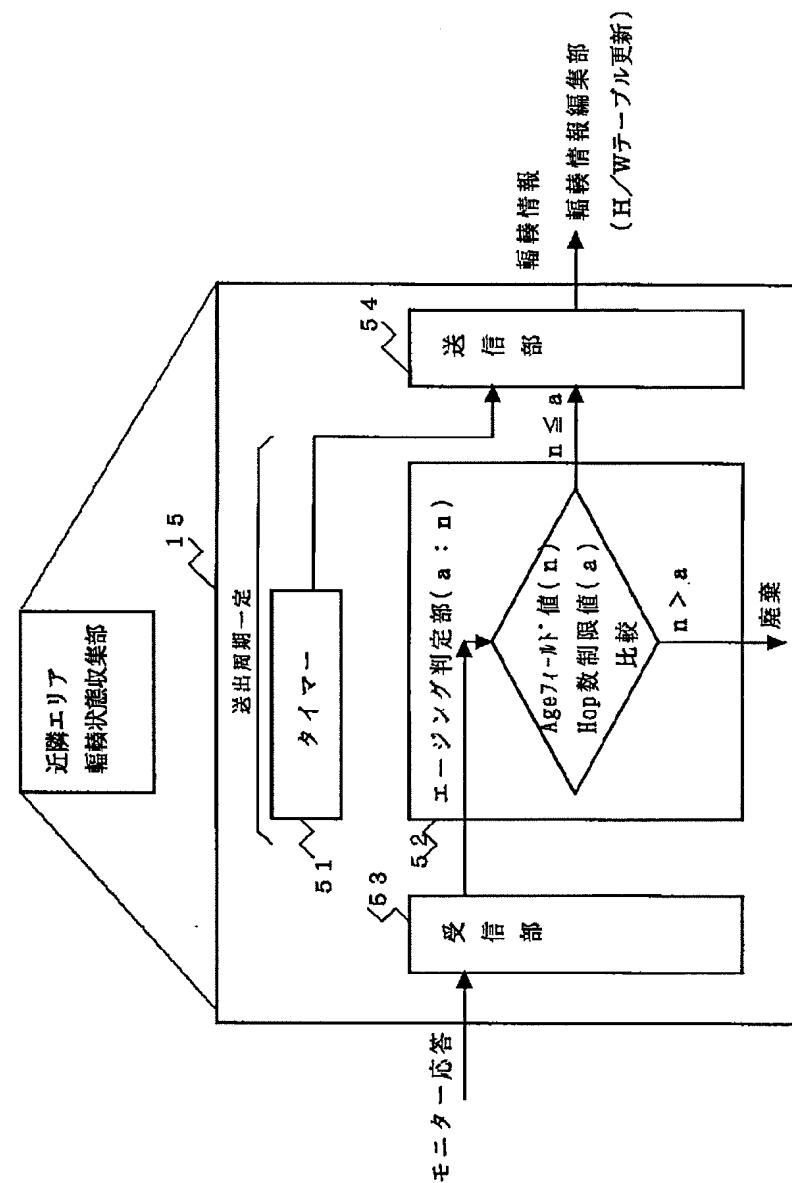
【図3】



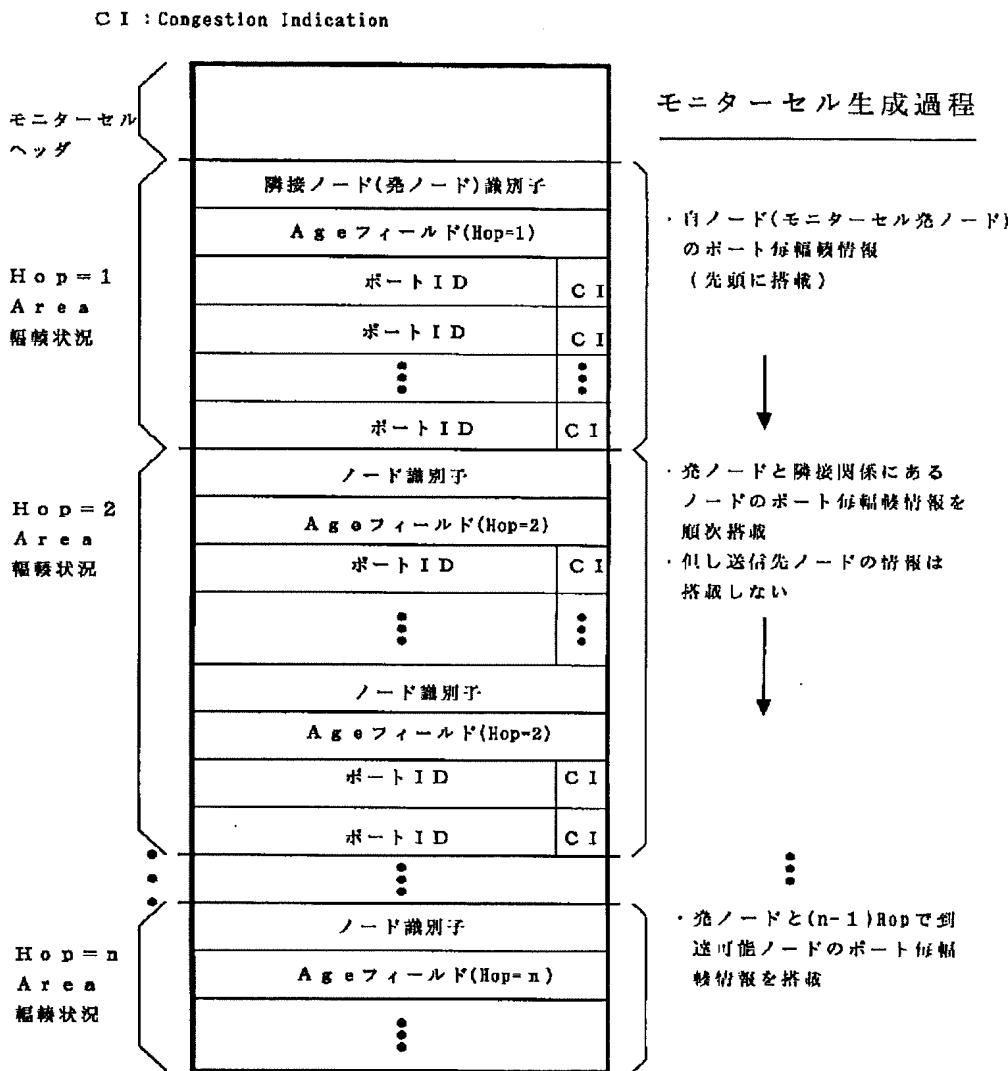
【図4】



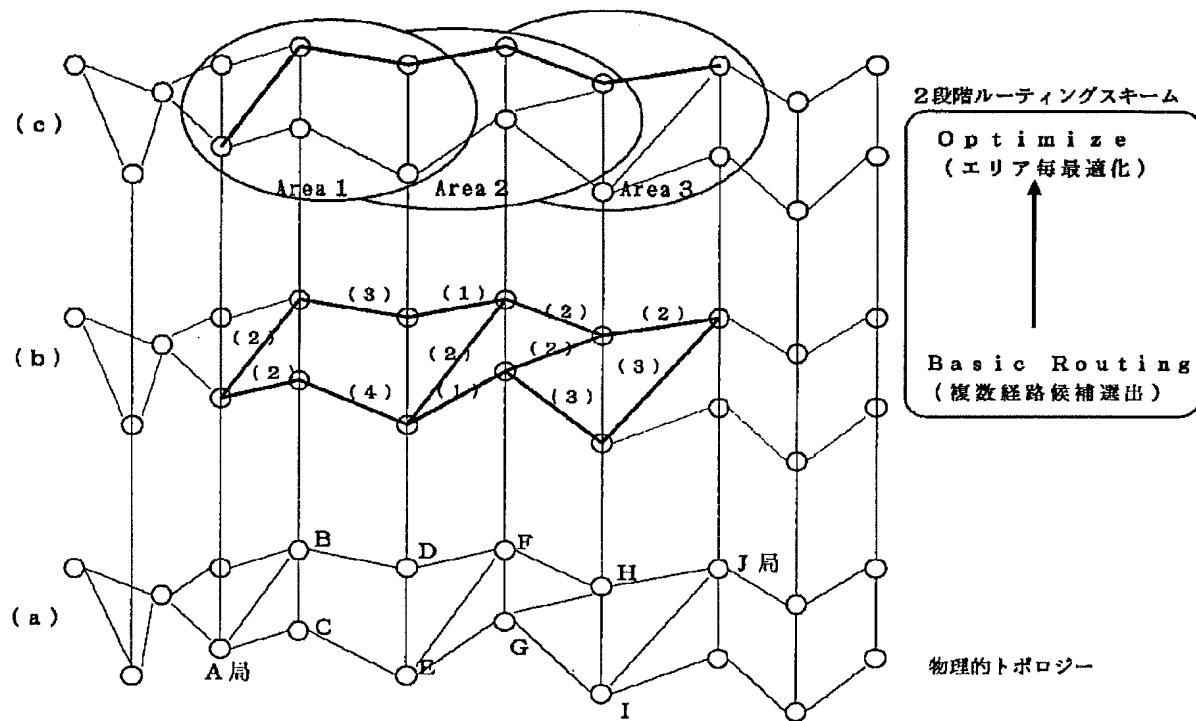
【図5】



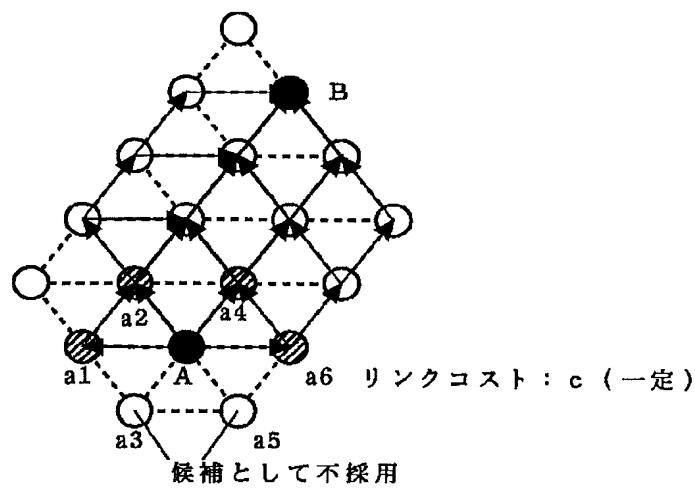
【図6】



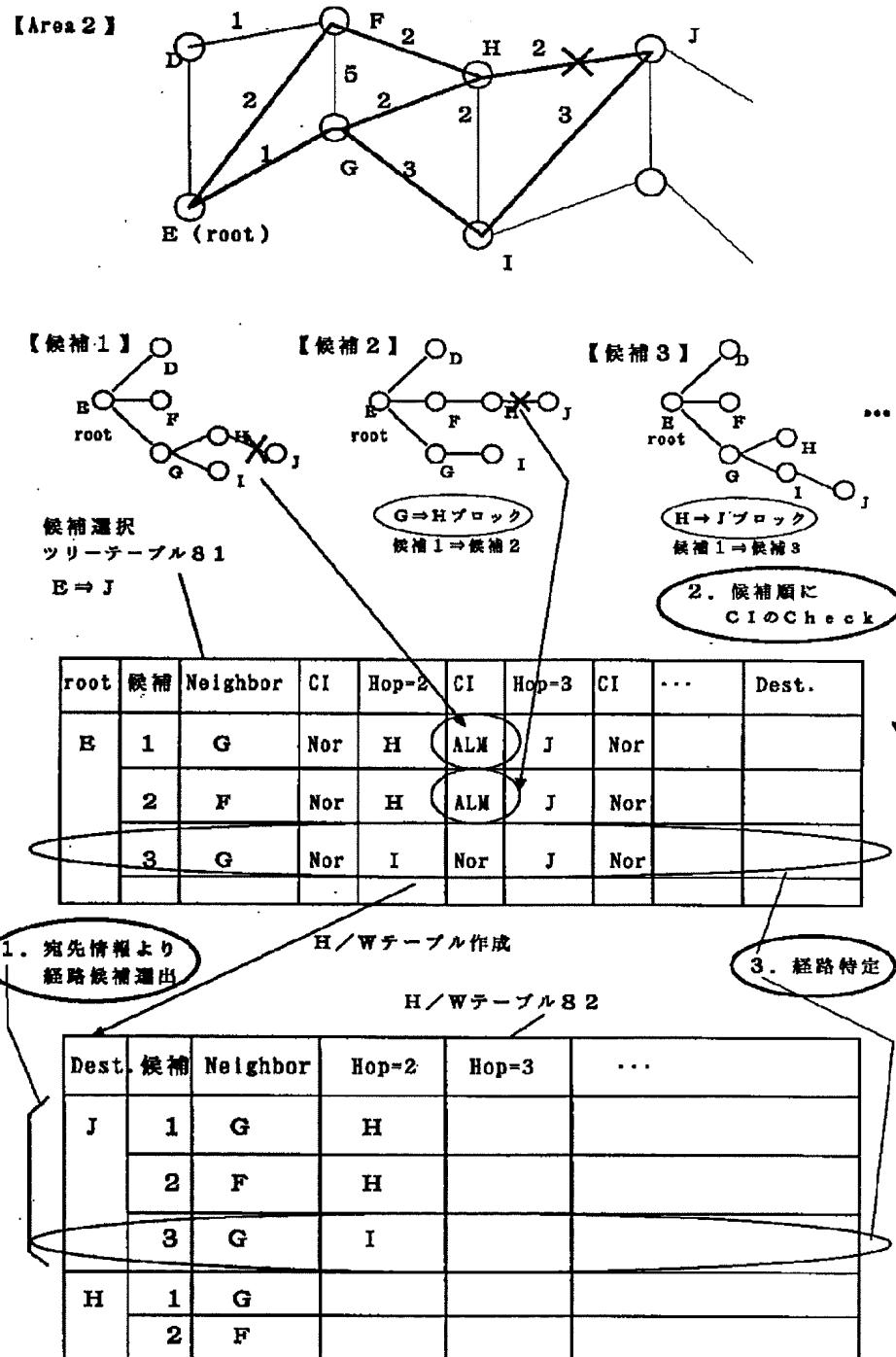
【図7】



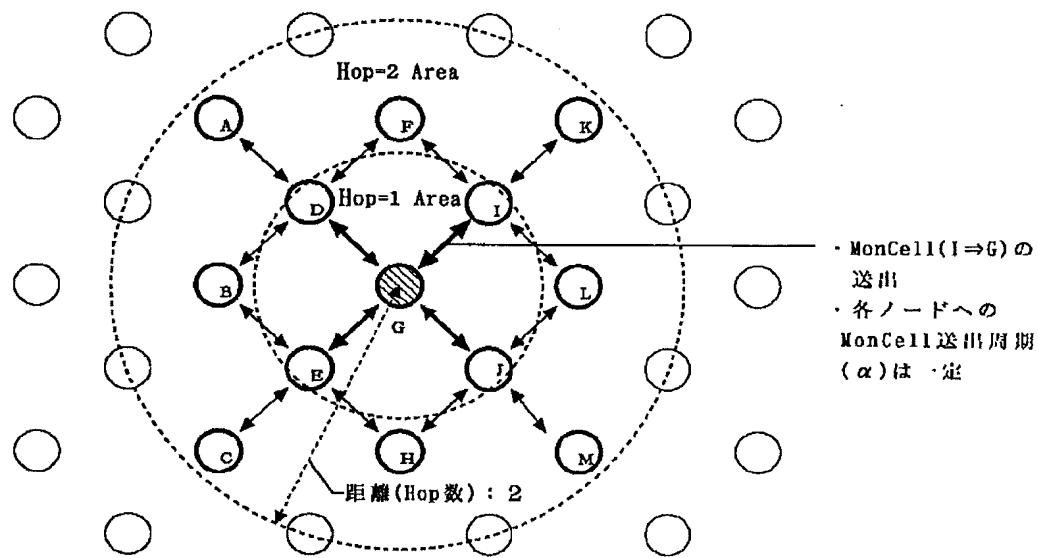
【図9】



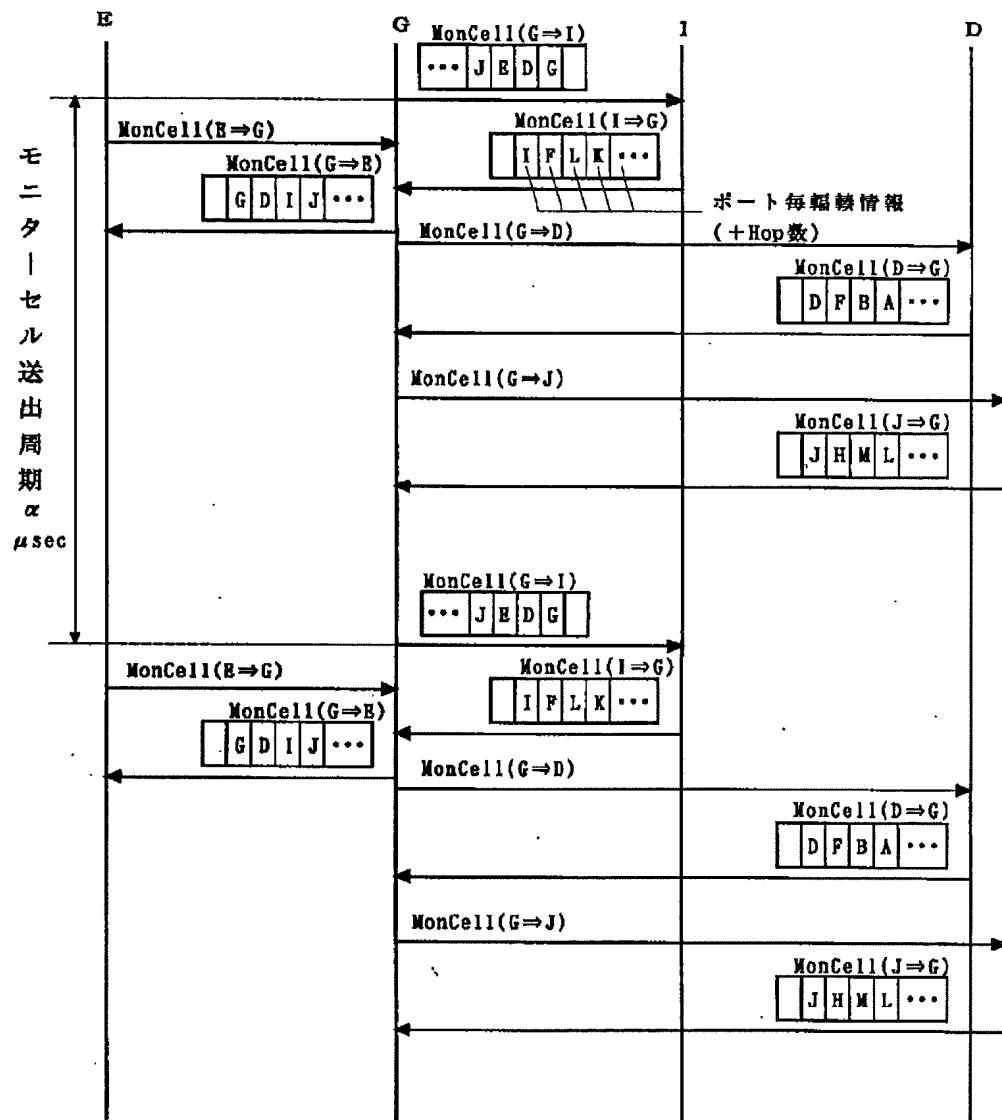
【図8】



【図10】



【図11】



## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-065733

(43)Date of publication of application : 06.03.1998

---

(51)Int.Cl. H04L 12/56

H04L 12/28

H04M 3/00

H04Q 3/00

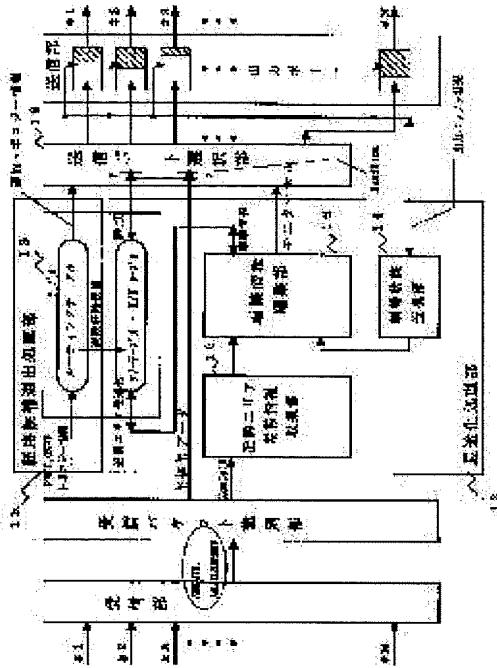
---

(21)Application number : 08-215678 (71)Applicant : NEC CORP

(22)Date of filing : 15.08.1996 (72)Inventor : MASUDA MICHIO  
NISHIHARA MOTOOS  
OGAWA AKIRA

---

(54) HIGH-SPEED ROUTING CONTROL SYSTEM



## (57)Abstract:

**PROBLEM TO BE SOLVED:** To provide a high-speed routing control system whereby optimum route selection is executed even when a destination node is adjacent, a change-over processing to a bypass circuit is executed at high speed, a traffic quantity by topology information is suppressed and follow-up is executed to the large scale conversion of a network.

**SOLUTION:** A route candidate selection processing part 11 selects plural route candidates till a destination node based on physical connecting information of a link which connects respective nodes and an optimizing

processing part 12 specifies the route candidate which does not include the link where a congestion occurs within the plural route candidates selected by the route candidate selection processing part 11 as an optimum route.

**LEGAL STATUS**

[Date of request for examination] 15.08.1996

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number] 2964957

[Date of registration] 13.08.1999

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

**\* NOTICES \***

JP0 and INPIT are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.\*\*\*\* shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

---

**CLAIMS**

---

**[Claim(s)]**

[Claim 1] In the routing control method which chooses the packet transfer path in the case of transmitting the packet which consists of connectionless data, such as LAN, through the ATM network of a connection mode to a predetermined destination node The path candidate election processing section which makes the multiple selection of the path candidate to a destination node based on the physical initial entry of the link which connects between each node, The high-speed routing control method characterized by having the optimization processing section which specifies the path candidate who does not include the link which congestion has generated among two or more path candidates elected by the path candidate election processing section as an optimal path based on the congestion condition of each link belonging to predetermined neighboring area.

[Claim 2] In a high-speed routing control method according to claim 1 the path candidate election processing section While total of the link cost value of the link which constitutes each path among the paths to a destination node elects the smallest minimum cost path based on the link cost value which shows the election priority of each link set up beforehand The inside of the path to the destination node which goes via each adjacent node which adjoins a self-node, The high-speed routing control method characterized by total of the link cost value of the path electing a path smaller than the total of the link cost value of said minimum cost path, and twice as many sum as the link cost value from a self-node to said adjacent node.

[Claim 3] In a high-speed routing control method according to claim 1 the optimization processing section While discarding the congestion information on the link which the node separated from the self-node more than the number of predetermined links among the congestion information which receives a predetermined monitor cel from

an adjacent node, and is stored in the monitor cel holds The high-speed routing control method which acquires the congestion information on each link which each remaining nodes hold, uses for specification of an optimal path, and is characterized by generating a new monitor cel from the acquired congestion information and the congestion information on each link which a self-node holds, and making it transmit to an adjacent node.

[Claim 4] It is the high-speed routing control method characterized by the optimization processing section transmitting a monitor cel for every predetermined period in a high-speed routing control method according to claim 3.

---

## DETAILED DESCRIPTION

---

### [Detailed Description of the Invention]

[0001]

[Field of the Invention] About a high-speed routing control method, especially this invention is an ATM network, by VPI and VCI of an ATM layer, identifies an IP packet only for the information on an ATM layer, and relates to the high-speed routing control method of the ATM connectionless communication device which performs transfer processing without packet assembly processing delay.

[0002]

[Description of the Prior Art] In recent years, the demand which holds the local connectionless property by the side of [ LAN ] a user (Local Area Network) with a global ATM network, and employs it by the high speed and the low price is increasing. In the original ATM network, if a connection setup is performed in advance of the main signal data transfer, each switch performs detailed cost evaluation, CAC (Connenction Admission Control) count, etc. for every connection setup and the optimal path is chosen from two or more candidates by signaling etc. simultaneously, the resource reservation for filling the QOS (Qualityof Service) demand from a user is performed.

[0003] In the routing protocol represented with PNNI (Private Network-to-Network Interface) in an ATM network, and OSPF (Open Shortest Path First) in IP communication link, in order to choose the optimal path, the "Dijkstra" algorithm is used. in order for a "Dijkstra" algorithm to calculate the minimum pass of the same cost based on the link information from all nodes, create the tree (root map: the minimum tree) to each destination which make itself the root (root: root) and to

perform a "Dijkstra" algorithm, each node manage (the link condition, i.e., the topology, which connect soon as path evaluation), and it be necessary to turn and notify the information to all nodes.

[0004] So, PNNI and OSPF have prescribed the technique called "Flooding" as an approach of notifying a link information. It is used in order that this approach may spread a packet including topology information immediately in fixed area, and actuation of supplying surrounding everybody widely if the received information is newer than the information which he has, and discarding if old has been to the base. Drawing 2 is the conceptual diagram having shown the distribution approach of the topology information by "Flooding" in PNNI. A. When 1-X.1 link condition changes, it is the propagation method characterized by for the node A.1 which detected the change serving as a source of dispatch, and advertising like a "flood" in the specific logical area PG (A).

[0005] On the other hand, there is no QOS demand from a user in the communication mode of the usual connectionless network represented by IP packet communication, and-like secondary optimization processing that a network side estimates a path by signaling processing is not performed. Conventionally, this kind of connectionless communication mode sets up the alternate route in addition to the direct path for the purpose of disconnection and the momentary thing it can be [ a thing ] satisfied with the maximum of a communication link demand, without on the other hand discarding a packet also to the concentration to a way, and the technique whose change processing to an alternate route is attained dynamically is used.

[0006] When an ATM network performs a connectionless communication link, two or more candidates who become two or more alternate routes as a method possessing the optional feature to detour pass are prepared, a surplus band value is supervised for every candidate path, and the method of performing control changed to the larger path of a surplus band is proposed (for example, JP,3-242063,A etc.). Drawing 3 is the block diagram showing an example of the conventional connectionless processor in an ATM network, and consists of a receive section 31, the partner address recognition section 32, the storage table section 33, the transmission-control section 34, and the transmitting section 35.

[0007] Drawing 4 is the flow chart which showed the flow of the conventional transmission-control section of operation. As shown in this drawing, if a connectionless packet is received, the conventional approach will extract a destination address from the additional information of a packet, and will choose PVC (Permanent Virtual Connection) corresponding to a destination address. As an

alternate-route setting means, the amount of PVC used of a direct path and two or more alternate routes is managed in the table storage section 33 of drawing 3, and it is the transmission-control section 34 and is going to satisfy a communication link demand of a user by choosing PVC which has an opening most as an alternate route. [0008]

[Problem(s) to be Solved by the Invention] However, there were the following troubles by such conventional routing control method. The 1st trouble is not accelerable [routing processing]. That is, since the traffic volume which routing information interchange takes when decision of the path by path evaluation of the link condition mold represented by the "Dijkstra" algorithm uses "Flooding" of a broadcasting mold as a propagation method increases, frequent information interchange is not performed. Therefore, a limitation is in real time nature and the optimality at the time of routing becomes coarse.

[0009] As the scrutinization means, by PNNI signaling in an ATM network, path evaluation with which are satisfied of QOS which it is detailed in 2nd order, and a user notifies by CAC processing, crank back processing, etc. for every connection setup is performed, and a routing path is decided. However, delay occurs by the overhead of this signaling processing. By the conventional alternate-route selection approach, in order to raise the precision of the permission band value of each path, a traffic fluctuation property is predicted, complicated count is needed in quest of a cost function faithful to real link topology fluctuation etc., and the judgment processing for changing to an alternate route cannot be accelerated. The factor which poses a problem has any approach in the high speed for drawing the alternate route used as an optimum solution and the processing means which gave priority to low delay not being established in convention time amount.

[0010] The 2nd trouble is that traffic volume increases in topology information propagation. the permission band of each path made into the decision criterion of the conventional cost evaluation -- the buffer pulse duty factor of each connectionless processor -- if it forces, depending on a traffic property, it changes every moment. Therefore, length face \*\*\*\* and routing table are updated by the fluctuation property of the permission band of each path, and the path control in a network lapses into a upset condition. Moreover, the traffic volume for advertising topology fluctuation to each connectionless communication device (ADOBATAIZU) increases. The factor which poses a problem is to use [ reflecting the surplus band value in cost information, and ] "Flooding" as a propagation method.

[0011] The 3rd trouble is being hard to follow in footsteps of network

large-scale-ization. If the direct factor which poses a problem originates in a propagation method being the "Flooding" mold like the above-mentioned and the number of connectionless communication devices is set to N, topology information propagation traffic volume will increase it to the order of O (N<sup>2</sup>). The three following points are considered as a means to control traffic volume. The 1st means is the approach of arranging the timer which controls traffic volume with each connectionless communication device. If the life (Lifetime) of the header unit of the packet which performs topology information interchange is used for the 2nd means and it becomes convention time amount, it will be the approach of annihilating a packet.

[0012] The 3rd means carries out division hierarchization of the logic area which carries out "Flooding", the advertising range is limited, and it is the approach of transmitting the summarized information, between between area and low order area, and high order area. The "Flooding" approach in the network hierarchized by drawing 2 is shown. Although the link, A.1-X.1 link and A.2-X.2 link, of two exists between PG(A)-PGes (X) of a lower layer, the upper layer shows carrying out "Flooding" of the topology information (Topology Aggregation) which collected the link of two to the link of one between A-X.

[0013] Due to such a control means, although too much traffic increase phenomenon can be eased to some extent, when the number of division area is set to m, topology information propagation traffic volume is still in an increase inclination to the order of O (N<sup>2</sup>/m), and cannot follow in footsteps of network large-scale-ization easily. Moreover, since an error arose at the time of topology information concentration when the number of division area (m) is increased and the advertising range of topology information is reduced, there was a trouble of losing the accuracy of routing as a result. Even when it approaches, while this invention's being for solving such a technical problem, and being able to carry out optimal routing and being able to accelerate the change processing to a detour, the amount of traffic using topology information can be controlled, and it aims at offering the high-speed routing control method which can follow network large-scale-ization.

[0014]

[Means for Solving the Problem] In order to attain such a purpose, the high-speed routing control method by this invention The path candidate election processing section which makes the multiple selection of the path candidate to a destination node based on the physical initial entry of the link which connects between each node, Based on the congestion condition of each link belonging to predetermined

neighboring area, it has the optimization processing section which specifies the path candidate who does not include the link which congestion has generated among two or more path candidates elected by the path candidate election processing section as an optimal path. Therefore, based on the physical initial entry of the link which connects between each node, the multiple selection of the path candidate to a destination node is made, and the path candidate who does not include the link which congestion has generated is specified as an optimal path based on the congestion condition of each link which belongs to predetermined neighboring area among these path candidates. Moreover, the path candidate election processing section is based on the link cost value which shows the election priority of each link set up beforehand. While total of the link cost value of the link which constitutes each path among the paths to a destination node elects the smallest minimum cost path. The inside of the path to the destination node which goes via each adjacent node which adjoins a self-node, Total of the link cost value of the path elects a path smaller than the total of the link cost value of the minimum cost path, and twice as many sum as the link cost value from a self-node to an adjacent node. Therefore, while the minimum cost path in which total of a link cost value is the smallest is elected among the paths to a destination node based on the link cost value of each link set up beforehand, the path among each path in which total of a link cost value is smaller than the total of the link cost value of the minimum cost path and twice as many sum as the link cost value from a self-node to an adjacent node is elected.

[0015] Moreover, while discarding the congestion information on the link which the node separated from the self-node more than the number of predetermined links among the congestion information which the optimization processing section receives a predetermined monitor cel from an adjacent node, and is stored in the monitor cel holds. The congestion information on each link which each remaining nodes hold is acquired, and it uses for specification of an optimal path, and a new monitor cel is generated from the acquired congestion information and the congestion information on each link which a self-node holds, and it is made to transmit to an adjacent node. Therefore, while the congestion information on the link which the node separated from the self-node more than the number of predetermined links among the congestion information stored in the received monitor cel holds is discarded. The congestion information on each link which each remaining nodes hold is acquired, and it is used for specification of an optimal path, and a new monitor cel is generated from the acquired congestion information and the congestion information on each link which a self-node holds, and it is transmitted to an adjacent node. Moreover, the optimization

processing section transmits a monitor cel for every predetermined period. Therefore, a monitor cel is transmitted for every predetermined period.

[0016]

[Embodyment of the Invention] Next, this invention is explained with reference to a drawing. Drawing 1 is the block diagram showing the high-speed routing control method of the connectionless communication device which is the gestalt of 1 operation of this invention, and it is characterized by being constituted by two steps of routing control methods which consist of the basic routing processing (Basic Routing) section 11 and the optimization processing (Optimize) section 12. The path candidate election processing section 11 has routing table 11a to which the physical initial entry of each node is set beforehand, and has the function to elect two or more global routing path candidates, based on the path evaluation algorithm of the link state mold used by the routing protocol represented with PNNI (Private Network-to-Network Interface) used with an ATM network, and OSPF (Open Shortest Path First) in IP communication link.

[0017] In two or more routing path candidates elected in the path candidate election processing section 11, the optimization processing section 12 performs still more detailed path evaluation, has the function to elect an optimal path as real time in the specific near area, and as shown in drawing 1, it consists of the table section 13, the congestion house keeping section 14, the neighboring area congestion information gathering section 15, and the congestion information editorial department 16. In the congestion house keeping section 14, own queue length or the own queue rate of increase of a buffer etc. is supervised, and it detects whether it is in a congestion condition per output port.

[0018] The neighboring area information gathering section 15 consists of the aging judging sections 52 which consider, the distance (Hop number), i.e., number of course links, of the sending-out section 54 which sends out a congestion status monitor cel, the receive section 53 which receives a congestion status monitor cel, and the received congestion status monitor cel, and perform an aging (Aging: life) judging, as shown in drawing 5. Moreover, the neighboring area information gathering section 15 possesses the timer 51, and has the function which generates a congestion status monitor cel periodically and sends it out by self-timing to the adjoining connectionless communication device.

[0019] Therefore, when the monitor cel which stored the congestion situation from the adjoining connectionless communication device is received in a receive section 53 and change is in the congestion situation, the neighboring area information gathering

section 15 searches the path candidate who removed the link with an adjoining connectionless communication device from the tree table of the table 13 mentioned later, and elects a semi-optimal path. Furthermore, with reference to a tree table, it reflects that the path candidate was changed in a H/W table.

[0020] The field which shows the Hop number of the connectionless communication devices via which it will have gone by the time it arrives is established in the monitor cel which each connectionless communication device collects with the information which shows the congestion situation of a buffer per output port of a neighboring connectionless communication device, and when the Hop number of the connectionless communication devices via which it goes becomes the count of a convention, it is used for the aging judging which controls an advertisement out of neighboring area by discarding applicable information. In the aging judging section 52, the need of he gaining and distributing to other adjacent nodes as compared with a Hop number of limiting value (a) to manage is judged, and either maintenance or abandonment is chosen and processed.

[0021] In the congestion information editorial department 16, it has the function to edit the monitor cel to distribute, with reference to the tree table of the own table 13. Drawing 6 is a monitor cel format, as shown in this drawing, describes an own congestion situation at the head, and generates a monitor cel in the format which inserted the information after the aging judging processing in the neighboring area information gathering section continuously after that. The point that congestion information is described according to the distance (Hop number) from a connectionless communication device is the description.

[0022] Next, actuation of this invention is explained with reference to drawing 7. Drawing 7 is the conceptual diagram of operation showing the high-speed routing control method which applied this invention, in (a), the physical topology of a connectionless communication device and (b) show actuation of the path candidate election processing section 11, and (c) shows actuation of the optimization processing section 12. Below, actuation of this invention is explained by making routing control actuation in the case of communicating in [ A ] -J games into an example. In addition, the link cost value which the inverse number of each link rate, delay characteristics, or a waste ratio is evaluated, and shows the election priority of each link at the time of network configuration shall be beforehand assigned to each link. The figure in () of drawing 7 (b) shows the link cost value of each link.

[0023] First, the path candidate election processing section 11 of each connectionless communication device acquires the link cost value of each link with

reference to routing table 11a, and the "Dijkstra" algorithm is used for it based on these links cost value, and it creates the root map (the minimum tree) to the J destinations which made themselves the root (root: root). [ A ] Furthermore, two or more paths which serve as a candidate for an alternate route are computed. With two or more paths here, total and the equivalent of the link cost value of the minimum cost path, or the conditions used as "Loop free" mentioned later are considered and elected. Two or more paths between the A-J offices elected by this processing are shown to drawing 7 (b) by the continuous line.

[0024] In the optimization processing section 12, in two or more candidate paths shown from the path candidate election processing section 11, optimization which considered the congestion situation in its near area is performed, and a routing path is specified. The optimal path elected by this processing is shown to drawing 7 (c) by the continuous line. Here, with reference to drawing 9, two or more candidate election process in path candidate selection section 11a is explained. Drawing 9 is the explanatory view showing two or more path election process.

[0025] Two or more paths which serve as a candidate must be paths which it is chosen in the basic routing section 11, and a duplication path (Looping), i.e., looping, does not generate. If total of the link cost value of each link from A games to a B station is set to NA and all node link costs are now set to c (fixed), total of the minimum link cost value will be set to  $NA=4c$ . Moreover, A games are adjoined, a1-a6 game exists, and the minimum link cost value between A-an ( $n:1-6$ ) is c.

[0026] It is  $Nan < NA+c=5c$  when the conditions which do not return to A games are taken into consideration, in order to set the link cost value between an-B to Nan ( $n:1-6$ ), and not to carry out looping, even if it goes via an, when transmitting to B from A. (formula 1)

It becomes. At this time, the link cost value to the B station of an game course found from A games is  $Nan+c < NA+c+c=6c$  (formula 2).

It becomes.

[0027] Therefore, the candidate set as the object of two or more paths is two kinds (the path to which the minimum link cost value between A-B which goes via the station becomes smaller than  $N+2c$ , i.e., "loop free"), when the minimum link cost value to the station adjoining, the path, i.e., minimum cost path, with the same minimum link cost value N, is set to c. In the example of drawing 9, the node which can be chosen from A games as a degree in two or more candidate selection process between A-B is limited to four adjacent nodes, a1, a2, a4, and a6. Since a link cost value is set to "6c", as for the path to which A games reach a B station via a3 or a5,

min also shows that it is excepted from the path candidate between A-B from (a formula 2).

[0028] Moreover, drawing 8 shows the process in which the path candidate elected in the basic routing section 11 is formed into a tree table, and the process in which the H/W table used for actual routing specifies a semi-optimal path with reference to a tree table in the transmit-port selection section 18. For example, when its attention is paid to E games which belong in Area2 of drawing 7 (c), the tree table 81 created manages the optimal path which makes the destination all other nodes that made themselves (E) the root per output port with CI (Congestion Indication : congestion display) flag.

[0029] The example of drawing 8 shows having elected the acquired path as "a candidate 2" and "a candidate 3", respectively, as a result of excepting a G-H link and a H-J link, respectively (block) and performing the "Dijkstra" operation in addition to the optimal pass of "a candidate 1." CI information is collected by the monitor cel, and specification of a semi-optimal path is reflected in the H/W table 82 after the check of a congestion situation is performed with reference to CI flag of the tree table 81.

[0030] Here, when it becomes clear that the buffer of the output port of a H-J link has started congestion from Node E among the paths to Node J, CI flag of the applicable part on the tree table 81 is updated by "ALM (congestion generating condition)" from "Nor (congestion discharge condition)." Therefore, the path of G-I-J of "a candidate 3" is chosen at the time of actual routing. In addition, about the creation procedure of the H/W table 82, the destination address which corresponds to the location of the leaf (leaf: leaf) of the tree table 81 shown from the basic routing section 11 first is searched (step \*\*), and the next step destination port corresponding to this destination address is determined. In that case, the congestion situation of path information is checked in order of the candidate shown from the tree table 81 (step \*\*), and an optimal path is specified (step \*\*).

[0031] Therefore, in the transmit-port selection section 18 of drawing 1, the 1st optimal path is chosen with reference to the H/W table 82 at the time of the stationary which will not have risk of being full of a buffer in the connectionless communication device in area soon. On the other hand, if a neighboring connectionless communication device turns out that the local buffer overflow has occurred, the method way of an output of the arbitration (the 2- the n-th) which avoided the congestion generating path out of given two or more candidates will be pinpointed, and control temporarily detoured per packet will be performed.

[0032] Thus, based on the physical initial entry of the link which connects between

each node, the multiple selection of the path candidate to a destination node is made in the path candidate election processing section 11. It is based on the congestion condition of each link which belongs to neighboring area in the optimization processing section 12. Since the path candidate who does not include the link which congestion has generated among two or more path candidates elected by the path candidate election processing section 11 was specified as an optimal path. Since a path candidate is elected based on static information with little fluctuation at the time of path candidate selection, while being able to shorten the processing time which election of a path candidate takes. Since an optimal path is specified based on the existence of a congestion generating link among these path candidates, the processing time which specification of an optimal path takes can be shortened, and routing control can be processed at a high speed as a whole.

[0033] Moreover, while total of the link cost value of the link which constitutes each path among the paths to a destination node elects the smallest minimum cost path based on the link cost value of each link set up beforehand, the path candidate election processing section. The inside of the path to the destination node which goes via each adjacent node which adjoins a self-node. Since total of the link cost value of the path elected the path smaller than the total of the link cost value of the minimum cost path, and twice as many sum as the link cost value from a self-node to an adjacent node. It becomes possible to sort out the path which overlaps and goes via a self-node by comparatively easy processing in advance, and optimal routing can be carried out even when a destination node approaches comparatively.

[0034] Next, actuation of the optimization processing section 12 is explained with reference to drawing 10. Drawing 10 is the conceptual diagram of operation having shown the description of the congestion situation propagation method in the optimization processing section 12 which applied this invention. Its attention is paid to the node G which constitutes the network of a skeleton pattern as shown in this drawing. Node G sends out a congestion status monitor cel by self-timing to Node I first. The congestion situation of adjacent nodes D, E, and J that Node G has acquired the own present congestion situation at the head now is added and sent out to the congestion status monitor cel to Node I like the above-mentioned (refer to drawing 6) from Node G.

[0035] Furthermore, Node G sends out a congestion status monitor cel similarly to the nodes D, E, and J which are other adjacent nodes, and the advertising range of the congestion situation of Node G is closed focusing on Node G as a result in the area of distance a (this example number of Age limit hop = 2). By own timing, Node G and the

node I in adjacency add the congestion situation of the adjacent nodes F, L, and K which are carrying out current acquisition of the current self-congestion situation at the head, and transmit a monitor cel to Node G. The nodes D, E, and J which are other adjacent nodes operate similarly, and Node G acquires the congestion situation of the node which exists in the distance a area centering on self.

[0036] According to the congestion situation propagation method in the optimization processing section which applied this invention, it becomes :O (N) thing which congestion information propagation traffic volume increases in proportion to the increment in the number of connectionless communication devices, and the rate of the increment in traffic volume is kept constant.

[0037] Drawing 11 is the timing diagram which showed the congestion status monitor cel sending-out timing of the node G in the skeleton pattern network of drawing 10. It is shown that each node performs timing which sends out a congestion status monitor cel the fixed period alpha (microseconde). A period is performed synchronizing with the clock in each connectionless communication device. In this case, as mentioned above ( drawing 5 , 6), based on the congestion status monitor cels 111-113 received from Nodes I, D, and J, the congestion information on each path of the tree table 81 is updated by the neighboring area congestion condition collection section 15 to a certain periodic timing.

[0038] Then, based on these congestion information, the congestion status monitor cel 114 addressed to node E is generated, and it is transmitted to the following periodic timing. Therefore, the information in the notice cel of a congestion situation becomes old in proportion to the distance of the connectionless communication device used as the information source, and even place [ most distant ], it can acquire the information on the past for "(distance a) x period (alpha)." The congestion information over a contiguity node serves as recently, and it turns out that it is suitable as information to acquire.

[0039] Thus, a predetermined monitor cel is received from an adjacent node in the optimization processing section 12. While discarding the congestion information on the link which the node separated from the self-node more than the number of predetermined links among the congestion information stored in the monitor cel holds Acquire the congestion information on each link which each remaining nodes hold, use for specification of an optimal path, and since a new monitor cel is generated from the acquired congestion information and the congestion information on each link which a self-node holds and it was made to transmit to an adjacent node It becomes possible to limit the advertising range of a monitor cel in physical neighboring area, the rate of

the increment in traffic by the monitor cel for collecting congestion information becomes always fixed to the increment in the number of nodes, and it can respond also to network large-scale-ization.

[0040] Moreover, since the monitor cel was transmitted for every predetermined period, the cel delay fluctuation of a monitor cel can be controlled, and fluctuation of the buffer pulse duty factor of the link buffer which each node has can be reduced. Furthermore, even when a lot of monitor cels are accidentally transmitted from some nodes, from an adjoining node, it becomes what is transmitted a normal period and the flood of the monitor cel traffic to the whole network can be prevented.

[0041]

[Effect of the Invention] The path candidate selection processing section which performs routing with global this invention as explained above, The optimization processing section which specifies the optimal path in predetermined neighboring area is prepared. In the path candidate selection processing section, the multiple selection of the path candidate to a destination node is made based on the physical initial entry of the link which connects between each node. Since the path candidate who does not include the link which congestion has generated in the optimization processing section based on the congestion condition of each link which belongs to predetermined neighboring area among these path candidates was specified as an optimal path Since a path candidate is elected based on static information with little fluctuation at the time of path candidate selection, while being able to shorten the processing time which election of a path candidate takes Since an optimal path is specified based on the existence of a congestion generating link among these path candidates, the processing time which specification of an optimal path takes can be shortened, and routing control can be processed at a high speed as a whole.

[0042] Moreover, it is based on the link cost value which shows the election priority of each link set up beforehand in the path candidate election processing section. While total of the link cost value of the link which constitutes each path among the paths to a destination node elects the smallest minimum cost path The inside of the path to the destination node which goes via each adjacent node which adjoins a self-node, Since total of the link cost value of the path elected the path smaller than the total of the link cost value of the minimum cost path, and twice as many sum as the link cost value from a self-node to an adjacent node It becomes possible to sort out the path which overlaps and goes via a self-node by comparatively easy processing in advance, and optimal routing can be carried out even when a destination node approaches comparatively.

[0043] Moreover, in the optimization processing section 12, a predetermined monitor cel is received from an adjacent node. While discarding the congestion information on the link which the node separated from the self-node more than the number of predetermined links among the congestion information stored in the monitor cel holds Acquire the congestion information on each link which each remaining nodes hold, use for specification of an optimal path, and since a new monitor cel is generated from the acquired congestion information and the congestion information on each link which a self-node holds and it was made to transmit to an adjacent node It becomes possible to limit the advertising range of a monitor cel in physical neighboring area, the rate of the increment in traffic by the monitor cel for collecting congestion information becomes always fixed to the increment in the number of nodes, and it can respond also to network large-scale-ization. Moreover, since the monitor cel was transmitted for every predetermined period, the cel delay fluctuation of a monitor cel can be controlled, and fluctuation of the buffer pulse duty factor of the link buffer which each node has can be reduced. Furthermore, even when a lot of monitor cels are accidentally transmitted from some nodes, from an adjoining node, it becomes what is transmitted a normal period and the flood of the monitor cel traffic to the whole network can be prevented.

---

[Translation done.]